



Scalability of Robotic Controllers: Speech-Based Robotic Controller Evaluation

by Rodger A. Pettitt, Elizabeth S. Redden, and Christian B. Carstens

ARL-TR-4858

June 2009

NOTICES

Disclaimers

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.

Army Research Laboratory

Aberdeen Proving Ground, MD 21005-5425

ARL-TR-4858**June 2009**

Scalability of Robotic Controllers: Speech-Based Robotic Controller Evaluation

Rodger A. Pettitt, Elizabeth S. Redden, and Christian B. Carstens
Human Research and Engineering Directorate, ARL

| REPORT DOCUMENTATION PAGE | | | Form Approved OMB No. 0704-0188 | | |
|---|-----------------------------|--|---|--|---|
| Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS. | | | | | |
| 1. REPORT DATE (DD-MM-YYYY) June 2009 | | 2. REPORT TYPE Final | | 3. DATES COVERED (From - To) August 2008–September 2008 | |
| 4. TITLE AND SUBTITLE Scalability of Robotic Controllers: Speech-Based Robotic Controller Evaluation | | 5a. CONTRACT NUMBER | | | |
| | | 5b. GRANT NUMBER | | | |
| | | 5c. PROGRAM ELEMENT NUMBER | | | |
| 6. AUTHOR(S) Rodger A. Pettitt, Elizabeth S. Redden, and Christian B. Carstens | | 5d. PROJECT NUMBER 62716AH70 | | | |
| | | 5e. TASK NUMBER | | | |
| | | 5f. WORK UNIT NUMBER | | | |
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) U.S. Army Research Laboratory ATTN: AMSRD-ARL-HR-MW Aberdeen Proving Ground, MD 21005-5425 | | 8. PERFORMING ORGANIZATION REPORT NUMBER ARL-TR-4858 | | | |
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) | | 10. SPONSOR/MONITOR'S ACRONYM(S) | | | |
| | | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) | | | |
| 12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited. | | | | | |
| 13. SUPPLEMENTARY NOTES | | | | | |
| 14. ABSTRACT This study, which took place at Fort Benning, GA, focused on the feasibility of reducing robotic controller size by replacing some of the manual controls with speech-based controls. Eleven Soldiers from the Officers Candidate School served as participants. After training on the operation of the iRobot PackBot small unmanned ground vehicle (SUGV) system, each Soldier teleoperated the SUGV using two controller conditions; a combination of speech and manual control and manual control only. Soldiers were tasked to drive the robot and to perform operations such as surveillance using the robotic arm. Controller type and usability were evaluated based on objective performance data, data collector observations, and Soldier questionnaire responses. Workload for each controller was measured by having the Soldiers complete the National Aeronautics and Space Administration Task Load Index survey after using each controller type. Speech-based control exhibited the potential for benefits beyond controller size reduction. It decreased time and effort when performing multiple tasks simultaneously by allowing speech commands to be given for control of the robotic arm while at the same time maneuvering the robot using manual controls. The speech-based control system also has the potential to provide other benefits beyond those addressed in this study. | | | | | |
| 15. SUBJECT TERMS teleoperation, speech-based, control scalability, robotic control, control size, dismounted soldiers | | | | | |
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT UU | 18. NUMBER OF PAGES 54 | 19a. NAME OF RESPONSIBLE PERSON Rodger A. Pettitt |
| a. REPORT Unclassified | b. ABSTRACT Unclassified | c. THIS PAGE Unclassified | | | 19b. TELEPHONE NUMBER (Include area code) 706-545-9142 |

Contents

| | |
|---|-----------|
| List of Figures | v |
| List of Tables | vi |
| 1. Introduction | 1 |
| 1.1 Statement of the Problem | 1 |
| 1.2 Overview of Experiment | 2 |
| 1.3 Objective | 2 |
| 2. Method | 3 |
| 2.1 Participants | 3 |
| 2.1.1 Pretest Orientation and Volunteer Agreement | 3 |
| 2.1.2 Medical Review and Screening | 3 |
| 2.2 Instruments and Apparatus | 3 |
| 2.2.1 The iRobot PackBot Explosive Ordinance Disposal (EOD) Robot | 3 |
| 2.2.2 Controller..... | 3 |
| 2.2.3 Robotic Course | 7 |
| 2.2.4 National Aeronautics and Space Administration (NASA) Task Load Index (TLX) | 8 |
| 2.2.5 Questionnaires | 8 |
| 2.3 Procedures | 8 |
| 2.3.1 Soldier Orientation | 8 |
| 2.3.2 Training | 9 |
| 2.3.3 Robotic Driving Course..... | 9 |
| 2.3.4 End of Iteration Questionnaire Administration | 10 |
| 2.4 Experimental Design | 10 |
| 2.4.1 Independent Variable | 10 |
| 2.4.2 Dependent Variables | 10 |
| 2.5 Data Analysis | 10 |
| 3. Results | 10 |
| 3.1 Demographics..... | 10 |
| 3.2 Training | 11 |

| | | |
|-----------|--|-----------|
| 3.3 | Robotic Course Results | 11 |
| 3.4 | NASA TLX Results..... | 13 |
| 3.5 | Questionnaire Results..... | 14 |
| 4. | Conclusions | 15 |
| 5. | Discussion and Recommendations | 15 |
| 6. | References | 16 |
| | Appendix A. Medical Status Form | 19 |
| | Appendix B. Demographics | 21 |
| | Appendix C. Training | 25 |
| | Appendix D. Post Iteration | 31 |
| | Appendix E. End of Experiment | 37 |
| | Distribution List | 42 |

List of Figures

| | |
|--|----|
| Figure 1. The iRobot PackBot EOD robot..... | 4 |
| Figure 2. Arm low pose. | 5 |
| Figure 3. Arm high pose. | 5 |
| Figure 4. Hand-held controller..... | 6 |
| Figure 5. Mean times (seconds) to complete stations..... | 11 |
| Figure 6. NASA TLX weighted means..... | 13 |

List of Tables

| | |
|---|----|
| Table 1. Speech commands..... | 4 |
| Table 2. Commands and related speeds. | 6 |
| Table 3. Robotic course stations. | 9 |
| Table 4. Paired comparisons, mean times to complete stations..... | 11 |
| Table 5. Mean driving errors, all stations. | 12 |
| Table 6. Mean IEDs located..... | 12 |
| Table 7. Mean correct, cognitive tasks. | 13 |
| Table 8. Mean correct, NASA TLX total workload ratings. | 14 |

1. Introduction

Speech-based systems are being investigated for many different applications. Speech-based systems have been evaluated for data entry (Mitchard and Winkles 2002; Tsimhoni et al., 2004), use by the disabled (Summers, 1988), assistance during medical examinations (Bravo, 2005), searching the web on handheld devices (Chang et al., 2002), and unmanned aerial vehicle (UAV) control (Draper et al., 2003). Speech-based input is an intuitive form of system control that can free both cognitive and physical operator resources. While speech-based systems show potential for military applications to free the hands and eyes during data input, they are not without potential problems. A primary challenge is intelligibility as affected by factors such as stress, noise, and speech mannerisms. These conditions are stated very well by Pigeon et al. (2005):

Military operations are often conducted under conditions of stress induced by high workload, high emotional tension, G-forces, and other conditions commonly encountered in the battlefield. These conditions are known to affect human speech characteristics. As a result, battlefield operations put military personnel under stress, which in turn impacts on their speech and the communications they have with speech processing equipment (such as voice coders, automatic speech (word) and speaker recognition) and other military personnel.

Other potential problems include background noise, other personnel speaking in the vicinity, discrete word usage versus normal speech, conflicts with intra-team communications, and time required to train the system to the user's voice (Draper et al., 2003; Williamson et al., 1996).

While challenges remain significant, progress has been made at providing speaker-independent, real-time speech recognition and investigations into military applications for speech-based systems have shown significant promise (Draper et al., 2003; Turner and Carstens, 2007). For example, Draper et al. (2003) found that speech input was superior to manual input for operators performing in a simulated teleoperated UAV control station environment. Advantages included reductions in task completion times, increases in task accuracy, increases in flight/navigation measures, and increases in pilot ratings. This study builds upon the investigation by Draper et al. (2003) by investigating the affect of speech-based control on small unmanned ground vehicle (SUGV) teleoperation.

1.1 Statement of the Problem

This experiment was the second in a series of experiments to investigate the effect of scaling robotic controls and displays for use by dismounted Soldiers. The first controller experiment focused on scalability in terms of reducing a controller size by either reducing size of individual controls or by reducing the number of controls through multifunction mapping (Pettitt et al., 2008). In that study, Soldiers used each type controller to maneuver a SUGV through a driving

course that required them to perform various reconnaissance and driving tasks similar to tasks they would normally perform during tactical operations. Several tasks required the operator to control multiple robotic functions simultaneously (e.g., raising the control arm while tilting the camera head). The multifunction controller limited the operator's ability to perform simultaneous tasks while tele-operating the SUGV resulting in significantly slower course completion times. Furthermore, Soldiers reported that the multifunction controller was more difficult to learn and switching between modes was time consuming and confusing.

This study focused on the feasibility of reducing controller size by replacing some of the manual controls with speech-based controls. Speech-based control also exhibits the potential for reducing mission completion time by enhancing the operator's ability to multi-task. For example, the operator could drive the SUGV with the operator control unit (OCU) while using speech commands to pan and zoom a camera, or switch camera views.

1.2 Overview of Experiment

This study investigated the effect of a speech-based robotic controller on SUGV control. It took place at Fort Benning, GA, and 11 Soldiers from the Officer Candidate School (OCS) participated in the study. The study was originally planned using 30 Soldiers over a 5-day period but due to inclement weather and troop availability, the scope of the study was reduced.

After training on the operation of the iRobot PackBot SUGV system, each Soldier teleoperated the SUGV using two controller conditions; combination of speech and manual control and manual control only. Controller type and usability were evaluated based on objective performance data, data collector observations, and Soldier questionnaires. Data were also gathered to document information requirements for teleoperated control of robotic ground vehicles.

The speech-based controller that was used in this experiment was developed by Think-A-Move (TAM). It used speech commands from sound captured in the ear canal. The in-ear capture of speech should allow a significant reduction of noise from external sources to improve command recognition accuracy in highly noisy environments. The controller allowed an operator to control SUGVs using a combination of speech and manual control inputs. The speech-based robotic controller gave the operator the ability to use speech commands to access functions that might otherwise require navigating through a menu or operating a succession of switches on the OCU. The speech-based control function was disabled during the manual operation trials.

1.3 Objective

The objective was to assess the effect of speech-based control on the operation of SUGVs.

2. Method

2.1 Participants

Eleven Soldiers from the OCS participated in the study.

2.1.1 Pretest Orientation and Volunteer Agreement

The Soldiers were given an orientation on the purpose of the study and what their participation would involve. They were briefed on the objectives and procedures, as well as on the robot. They were also told how the results would be used and the benefits the military could expect from this investigation. Any questions the subjects had regarding the study were answered. It was made clear that Soldier participation in the experimentation was voluntary. The Volunteer Agreement Affidavit (VAA) was explained and its contents verbally presented. The Soldiers were given the VAA to read and sign if they decided to volunteer.

2.1.2 Medical Review and Screening

The Soldiers were given a medical status form (appendix A) to determine if any of them had a medical profile or history that would jeopardize them if they participated in the study. The medical review forms were reviewed by the principle investigators and none of the Soldiers indicated a medical condition that would preclude their participation.

2.2 Instruments and Apparatus

2.2.1 The iRobot PackBot Explosive Ordinance Disposal (EOD) Robot

The iRobot PackBot EOD Robot (figure 1) is a portable SUGV reconnaissance and tactical robot that can enter and secure areas that are either inaccessible or too dangerous for humans. The robot is equipped with an OmniReach manipulator system to allow it to extend over 2 m, when examining suspicious objects on EOD missions. The iRobot PackBot EOD Robot payload has a rotating pan and tilt head equipped with multiple cameras.

2.2.2 Controller

The TAM speech-based robotic controller uses speech commands from sound captured in the ear canal. The controller allows an operator to control unmanned ground vehicles (UGVs) using a combination of speech and manual control inputs. The system was configured so that concurrent control was possible, i.e., the operator could control separate functions on the robot using the hand-held controller and speech commands at the same time. For example, while driving the robot using the hand-held controller, the operator could switch cameras or zoom in/out using speech commands without having to stop the robot. Speech commands used during this study are shown in table 1.

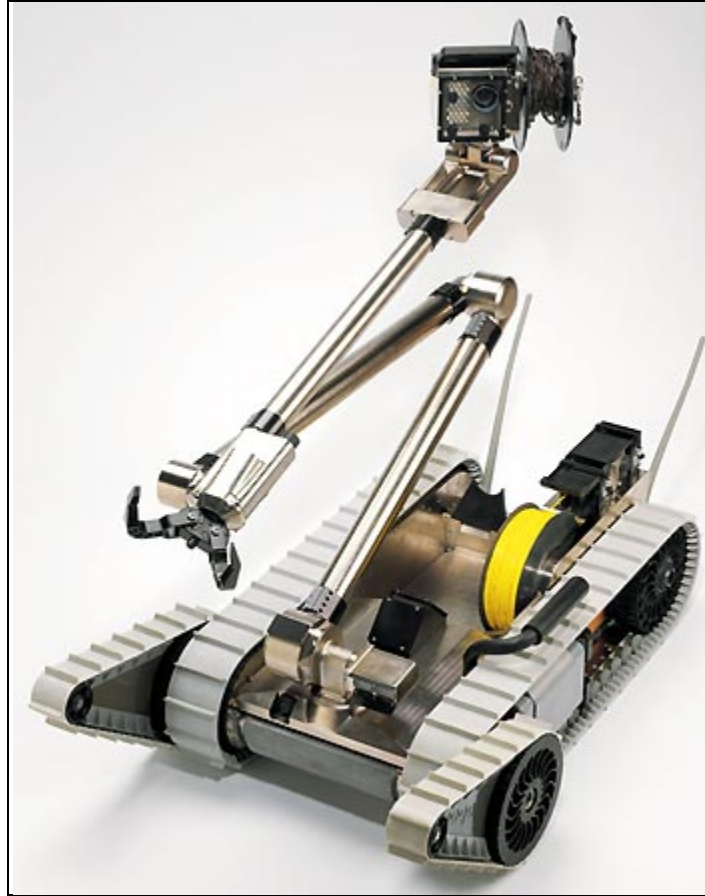


Figure 1. The iRobot PackBot EOD robot.

Table 1. Speech commands.

| Arm Poses | Camera Commands | Motion Commands | Other Commands |
|-------------|-----------------|-----------------|------------------|
| Arm drive | Pan left | Halt | Flippers forward |
| Arm deploy | Pan right | Forward small | Flippers back |
| Arm retract | Tilt up | Reverse small | Light on |
| Arm high | Tilt down | Left small | Light off |
| Arm low | — | Right small | — |
| — | — | Left large | — |
| — | — | Right large | — |
| — | — | Brake on | — |
| — | — | Brake off | — |

The arm pose commands were used to extend or retract the robotic arm into preset arm positions enabling the operator to quickly position the camera for observation. Figure 2 shows the arm in the arm low position and figure 3 shows it in the arm-high position. Manual activation of the arm poses was accomplished by using the touch-screen menu located on the display screen.



Figure 2. Arm low pose.



Figure 3. Arm high pose.

The camera panned or tilted a predetermined consistent distance each time a pan or tilt command was given. Small and large motion commands controlled the speed of the robot, i.e., when the “left small” command was given, the robot turned at a slow speed ($5^\circ/\text{s}$) and when the command “left large” was given, it drove forward at a faster speed ($45^\circ/\text{s}$). Table 2 shows the robot speed for each command.

Table 2. Commands and related speeds.

| Robot Motion Commands | Speed |
|-------------------------------------|----------|
| “Forward small” and “reverse small” | 0.25 m/s |
| “Left large” and “right large” | 45°/s |
| “Left small” and “right small” | 5°/s |
| Camera Control Commands | — |
| “Pan left” and “pan right” | 57°/s |
| “Tilt up” and “tilt down” | 57°/s |
| Flipper Control Commands | — |
| “Flippers forward” | 200°/s |
| “Flippers back” | 200°/s |

The system was designed so that the hand-held controller took precedence over speech commands for controlling the same functionality, i.e., if the robot was being driven forward using speech, activating the joystick transferred control to the hand-held controller. During the manual control iterations, the speech function was disabled, requiring the Soldier to input all commands manually. Figure 4 shows the hand-held controller used during this study.

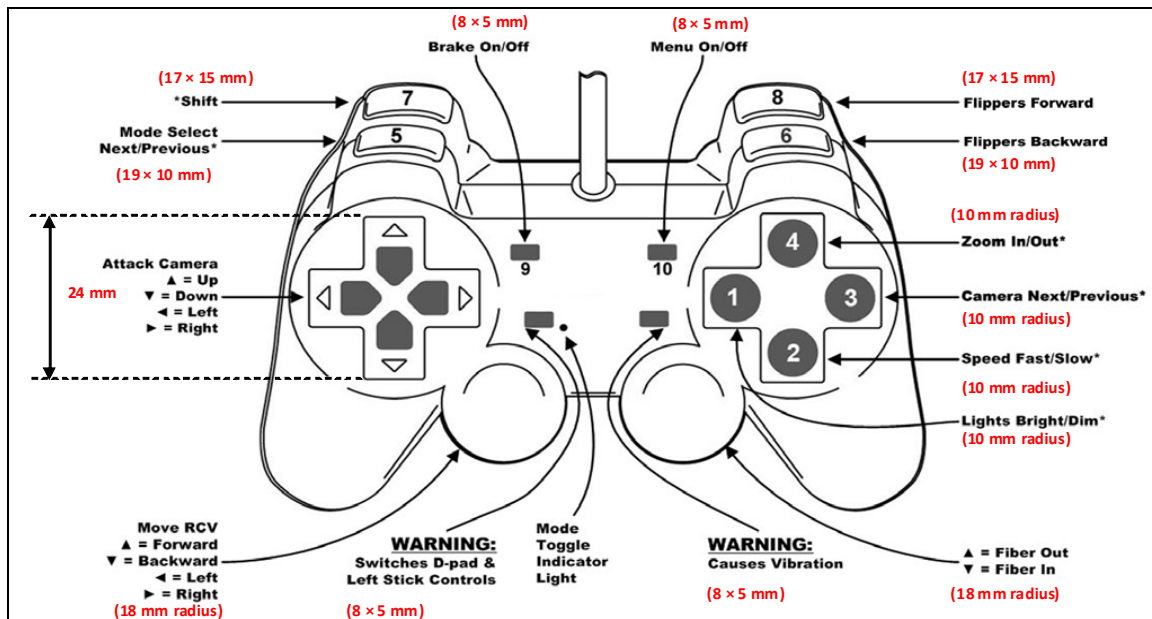


Figure 4. Hand-held controller.

For this study, Soldiers used a “push-to-command” button, which they depressed when they were issuing speech commands, and released after they completed issuing the command. The button was attached to the bottom of the hand-held controller.

2.2.3 Robotic Course

The experiment was conducted on a driving course ~150 m long and 1 m wide and clearly marked with white engineer tape on the left and right sides. Soldiers were located out of the line-of-site of the robot when maneuvering the robot through the course. Along the course, there were five stations where the operator conducted a specified maneuver or reconnaissance task. The five stations were window reconnaissance, target identification/tracking, area search, bunker reconnaissance, and tunnel reconnaissance. A description of each station and instructions for executing them follows.

- Station 1. Window Reconnaissance. The operator's task at this station is to locate and identify the simulated improvised explosive device (IED) hidden behind the urban wall. Upon reaching the wall, the operator must position the robot below the window while extending the sensor head arm and maneuver the sensor head in order to see through the window opening. The operator locates and identifies to the data collector a pipe bomb, ammo can bomb, or soda can bomb.
- Station 2. Target Identification/Tracking. The operator positions the robot behind a four foot wall while simultaneously extending the sensor head arm. The operator zooms in on the human target and identifies the object he is holding as a sledge hammer, M4 carbine, shovel, or box. The human target is then cued by the data collector to start walking perpendicular to the robot's line of sight. At set intervals along the route, the human target discards the object and replaces it with a different object. The operator is required to track the target and report to the data collector what the target is carrying at all times throughout the exercise.
- Station 3. Area Search. The search area is ~20 m long and 10 m wide. The operator maneuvers the robot through the area and searches for five simulated IEDs placed at various locations in the search area. The operator is given 5 min to locate and report the IEDs. During the area search, a data collector gives the operator a challenge and password, time of attack, and direction of attack. At the end of the area search, the operator is asked to recall the information he was given to evaluate cognitive workload. The data collector records the accuracy of information recalled by the operator.
- Station 4. Bunker Reconnaissance (Foxhole). The operator's task at this station is to locate and identify the simulated IED marked with an alphanumeric label hidden in a partially covered foxhole. Upon the robot reaching the foxhole, the operator must back the robot into position and extend the sensor head arm and maneuver the sensor head in order to see through the foxhole opening.
- Station 5. Tunnel Reconnaissance. The tunnel is a culvert pipe 6 m long \times 1 m in diameter and made of corrugated steel. A simulated IED is placed in the tunnel. The operator aligns the robot with the tunnel opening and maneuvers the robot through the pipe. Once the

operator locates the IED, he must use the camera zoom feature and report the alphanumeric label to the data collector.

2.2.4 National Aeronautics and Space Administration (NASA) Task Load Index (TLX)

The NASA TLX is a subjective workload assessment tool which allows subjective workload assessments on operator(s) working with various human-machine systems (Hart and Staveland, 1988). It uses a multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six subscales. These subscales include mental demands, physical demands, temporal demands, own performance, effort, and frustration. It can be used to assess workload in various human-machine environments such as: aircraft cockpits; command, control, and communication workstations; supervisory and process control environments; simulations; and laboratory tests. The version of the NASA TLX used during this experiment was presented to the Soldiers on a computer. Definitions of each scale were provided on laminated paper so the participants could refer to it as they were providing their estimates of the workload associated with each type of SUGV control on the various scales.

2.2.5 Questionnaires

The questionnaires were designed to elicit Soldiers' opinions about their performance and experiences with each of the controller conditions. The questionnaires were adapted from Pettitt et al. (2008) and modified for the specific details for the present study. The questionnaires asked the Soldiers to rate the devices on a 7-point semantic differential scale ranging from "extremely bad/difficult" to "extremely good/easy." Questionnaires were administered at the end of each iteration and at the end of the experiment. Training questionnaires designed to rate the amount of practice time given, the level of detail presented, and the adequacy of training were administered at the completion of training. Questionnaires were also used to gather information concerning the Soldiers' demographic data, robotic experience, and physical characteristics that might affect their ability to operate the robot (see appendices B–E).

2.3 Procedures

2.3.1 Soldier Orientation

Upon arrival, the experiment Soldiers received a roster number, which was used to identify them throughout the evaluation. After their pretest orientation, the VAA was explained and its contents verbally presented. Next, the Soldiers were given the VAA to read and sign if they decided to volunteer. The volunteers were then asked to complete the Medical Review and Screening form. Demographic data, as well as data concerning their U.S. Army and robotic experience and physical characteristics, were also taken for each Soldier. Next, Soldiers received an orientation on the robotics driving course.

It was explained to them that upon completion of training they would complete two iterations of the robotic driving course (one with each controller type). Half the participants drove the course using the combination of speech and manual control first, then again using manual control). The other half completed the course with manual control then again using the speech-based controller. After each iteration, Soldiers completed an end of iteration questionnaire designed to elicit their opinions about their performance and experiences with each of the controller conditions. After the second iteration, they completed an end of exercise questionnaire.

2.3.2 Training

A representative from TAM trained the Soldiers on the operation of the robot and controllers. Soldiers practiced teleoperating the robot on the actual course used during the experiment. The course required driving, obstacle negotiation, sensor head manipulation, target identification, and moving target tracking. It also involved training, on the graphic user interface (GUI) and use of the video cameras. During the training each Soldier practiced performing robotic tasks at each station using both speech-based and manual control manipulations. Upon completion of the training, the Soldiers were administered training questionnaires designed to rate the amount of practice time given, the level of detail presented, and the adequacy of training.

2.3.3 Robotic Driving Course

Soldiers negotiated the robotic driving course by teleoperating the robot using both controller conditions. The course required driving, obstacle negotiation, sensor head manipulation, target identification, and moving-target tracking.

The robotics course was ~150 m in length and had five stations along its route (see table 3). The Soldiers were told to maneuver the robot along the course marked with engineer tape and conduct a reconnaissance at key locations along the route. At each station, the operator was instructed to perform a specified maneuver or reconnaissance task. A data collector following the robot recorded the time required for the Soldier to complete each station and number of driving errors. A driving error was defined as running into objects on the course or failing to properly position the robot to perform the task required at a station. An additional data collector was co-located with the Soldier to observe and record the Soldier's responses and to record overall course completion times.

Table 3. Robotic course stations.

| Station | Description | Task |
|---------|---------------------------------|--|
| 1 | Window reconnaissance | Maneuver robot into position and locate IED |
| 2 | Area scan/moving target | Maneuver robot into position, locate, track and report target activities |
| 3 | Area search | Maneuver robot and search area for IEDs |
| 4 | Bunker reconnaissance (foxhole) | Maneuver robot into position and locate IED |
| 5 | Tunnel reconnaissance | Maneuver robot into position and locate and identify IED |

2.3.4 End of Iteration Questionnaire Administration

Questionnaires, designed to assess participants' performance and experiences with each control system, and the NASA TLX were administered to each Soldier at the end of each iteration. After completing the course with each controller condition, the Soldiers completed an end of experiment questionnaire that compared both controller conditions on a number of characteristics. They also completed questionnaires concerning the information requirements for teleoperating the robot.

2.4 Experimental Design

The design of this experiment was a single factor repeated measures design.

2.4.1 Independent Variable

- Controller condition, manual vs. speech based

2.4.2 Dependent Variables

- Time to complete each station on the course
- Number of driving errors
- Number of IEDs identified
- Number of correct cognitive tasks
- NASA TLX workload scores
- Questionnaire responses

2.5 Data Analysis

Performance data were analyzed using paired samples *t*-tests. Cohen's *d*, an index of effect size, was computed for each *t*-value. Cohen's *d* is the difference in means divided by the pooled standard deviation. By convention, a value of $d = 0.2$ is considered a small effect size, $d = 0.5$ is considered to be a medium effect size, and $d = 0.8$ is considered to be a large effect size.

Iteration effects were controlled through the counterbalanced order of the experimental design.

Soldier questionnaire data were analyzed using descriptive statistics on the subjective ratings.

3. Results

3.1 Demographics

The Soldiers ranged in rank from E4 to E6. None of the Soldiers had any prior experience in teleoperating a robot. Detailed responses to the demographics questionnaire are available in appendix A.

3.2 Training

The participants rated the training as being very good for both robotic control systems.

3.3 Robotic Course Results

Figure 5 shows the average times to complete each of the five stations. There was a trend for slightly faster mean times with the manual control system. A series of *t*-tests comparing the mean times at each station is summarized in table 4. Only one of the comparisons was statistically significant: manual control was significantly faster than speech-based control at the window station. Even though most of the *t*-test values were non-significant, the magnitude of the *d* values suggests that that statistical significance might have been obtained with a larger sample size.

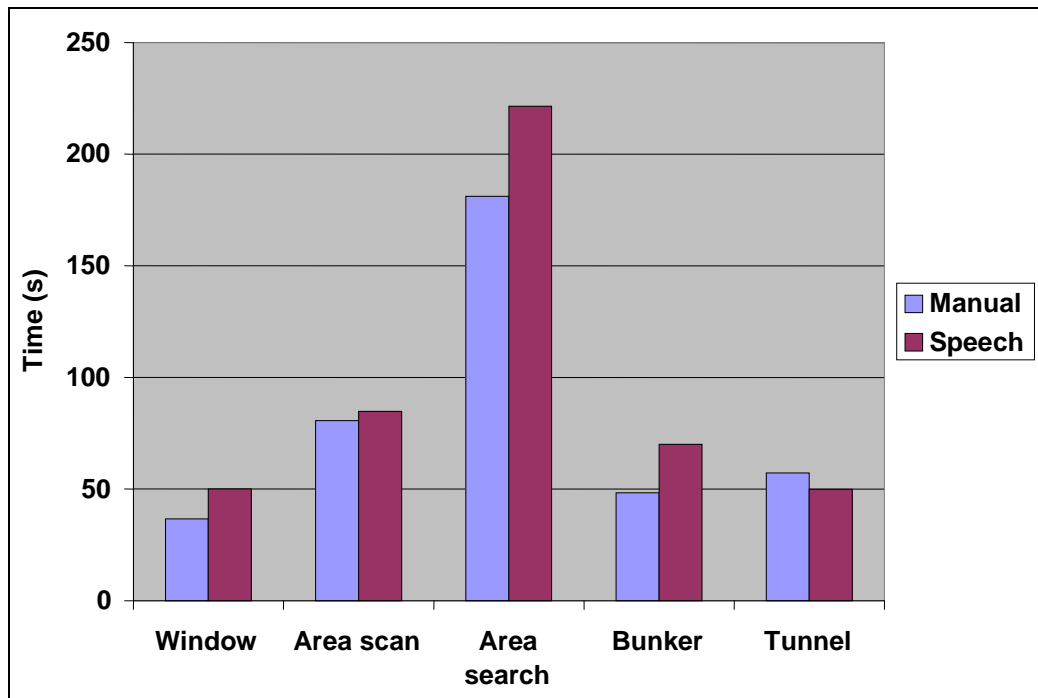


Figure 5. Mean times (seconds) to complete stations.

Table 4. Paired comparisons, mean times to complete stations.

| Station | t | df | p | d |
|-------------|-------|----|--------------------|------|
| Window | -2.82 | 10 | 0.018 ^a | 0.82 |
| Area scan | -0.75 | 10 | 0.469 | 0.30 |
| Area search | -1.48 | 10 | 0.170 | 0.59 |
| Bunker | -1.54 | 10 | 0.155 | 0.66 |
| Tunnel | 0.95 | 10 | 0.364 | 0.40 |

^a $p < 0.05$, 2-tailed.

At the window, area scan, bunker and tunnel stations the Soldiers were able to position the robotic arm faster with the voice controls than with the manual controls. The increased times at the window, area scan, and bunker positions using the speech-based commands were attributed to the incremental movement of the pan and tilt function of the camera. Soldiers had to make repeated voice commands to move the pan and tilt mechanism left, right, up, and down (each command only moved the pan and tilt a small distance; therefore a large movement required multiple repeated voice commands). This resulted in total station times that were greater for the voice controller. Changing the pan and tilt movement from a discrete movement to a continuous movement may decrease the time required for this movement. With continuous commands, a soldier could give one command to begin the desired action (e.g., camera pan or tilt) and another command to stop the action.

Table 5 shows the mean number of driving errors summed across all stations and the standard deviation (SD). There was no significant difference in the mean driving errors for the manual vs. speech-based control systems: $t(df = 10) = 0.45$, $p = 0.659$, and $d = 0.22$.

Table 5. Mean driving errors, all stations.

| Control | Mean | SD |
|--------------|------|------|
| Manual | 0.55 | 1.04 |
| Speech-based | 0.36 | 0.50 |

At the window and area scan stations, the participants had to simultaneously operate the robot arm while positioning the robot. The Soldiers were successful at this task 100% of the time with manual controls, while they were successful on 86% of the trials with the speech-based control system. On several occasions, the speech-based commands had to be repeated before the system responded to the command causing the arm to deploy after the robot was already in position to observe.

The Soldiers were directed to locate IEDs at the area search, bunker, and tunnel stations (table 6). There was no significant difference between manual and speech-based controls: $t(df = 10) = 1.75$, $p = 0.111$, and $d = 0.80$.

Table 6. Mean IEDs located.

| Control | Mean | SD |
|--------------|------|------|
| Manual | 6.91 | 0.30 |
| Speech-based | 6.36 | 0.92 |

The participants were given three cognitive tasks to complete at the area search station. The mean number of correct responses is shown in table 7. There was no significant difference between manual and speech-based controls: $t(df = 10) = 0.77$, $p = 0.459$, and $d = 0.33$.

Table 7. Mean correct, cognitive tasks.

| Control | Mean | SD |
|--------------|------|------|
| Manual | 1.73 | 0.90 |
| Speech-based | 1.36 | 1.36 |

Data collectors recorded whether the Soldiers were able to successfully position the robot at the window, area scan, and bunker stations. All attempts were successful with both the manual and speech-based control systems. Similarly, all attempts at locating targets at the window and area scan stations were successful, regardless of the control system used.

3.4 NASA TLX Results

The weighted means for the six NASA TLX scales are shown in figure 6. None of the differences between the speech-based and manual means were statistically significant. However, the trend toward more frustration with the speech-based control system approached statistical significance: $t(df = 10) = 1.98$, $p = 0.076$, and $d = 0.50$.

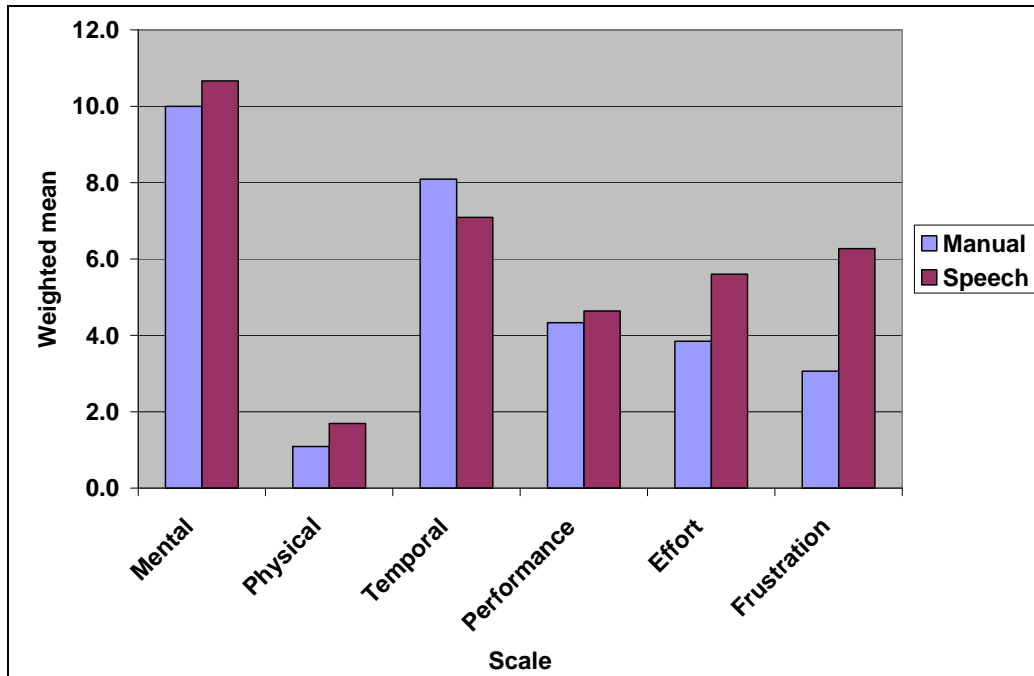


Figure 6. NASA TLX weighted means.

Soldiers that rated their frustration as high when using the voice control frequently stated that they were having difficulty getting the system to respond to their instructions the first time they were given and thus had to repeat the commands. Some of this difficulty could be attributed to this configuration of the TAM system which required the Soldiers to use a “Push-to-Command”

(PTC) button when issuing speech commands. The correct use of the button required the participants to depress the button before starting to speak the command and release it after finishing the command. The system was quite sensitive to the correct use of the PTC button, that is, the system was very likely to make errors in recognition if the button was depressed/released midway through the command. Data analysis on the eight participants for whom sound files were recorded showed that 66.5% (107 of 161) instances of mis- or non-recognition of the speech commands were due to incorrect use of the PTC button or the incorrect speech command being given. Elimination of the need to depress the PTC button for every speech command could remedy this problem. Speech recognition accuracy, excluding these instances, was 90.72% for these eight participants. Further analysis showed that for seven of the eight participants, speech recognition accuracy was 96.14%. Speech recognition accuracy for the eighth participant was 73.76%. Improvements to both speech processing algorithms and training procedures could increase speech recognition accuracy.

The mean total workload ratings for speech-based and manual controls are shown in table 8. The difference between means was not statistically significant: $t(df = 10) = 1.78$, $p = 0.106$, and $d = 0.29$.

Table 8. Mean correct, NASA TLX total workload ratings.

| Control | Mean | SD |
|--------------|------|------|
| Manual | 30.4 | 19.1 |
| Speech-based | 36.0 | 19.0 |

3.5 Questionnaire Results

The participants rated most of the tasks as being easier with the manual control than with the speech-based control, although the mean ratings differences were not large. The greatest preferences for manual vs. speech control occurred in response to items concerning discrete pan and tilt commands. When responses to the three tasks related to discrete pan and tilt commands were excluded, the average scores between manual control and speech control were essentially the same, with both command types receiving the rating of “very easy to use.” Changing these functions to continuous movement commands could increase the ratings for speech control of the pan and tilt functions. The participants rated performing simultaneous tasks, such as manipulating the robot arm and flippers using voice commands while manually driving the robot, as being easier with the speech-based control.

The participants rated the speech-based control system as being slightly better than the manual system in terms of control responsiveness, location of controls, and mapping of the control functions. Soldiers commented that it was easier to remember the voice commands than the locations of the controls on the hand held controller.

At the end of the experiment, seven of the participants expressed an overall preference for the manual controls system, while three preferred the speech-based system. The preference question was a forced choice question that required the Soldiers to pick one control system or the other, but several Soldiers commented that they would like a combination of voice and manual control. Even though the manual system was generally preferred, all of the participants commented that the speech-based controller shows potential for robotic operation, assuming that improvements could be made in the voice control system. Detailed responses to the post iteration questionnaire are available in appendix C and the end of experiment questionnaire in appendix D.

4. Conclusions

This study focused on the feasibility of reducing controller size by replacing some of the manual controls with speech-based controls. In this study, speech-based control exhibited the potential for benefits beyond controller size reduction. It decreased time and effort when performing multiple tasks simultaneously by allowing speech commands to be given for control of the robotic arm while at the same maneuvering the robot using manual controls. The speech-based control system also has the potential to provide other benefits beyond those addressed in this study. Certain tasks, such as menu navigation can be extremely time consuming and detrimental to the Soldier's situation awareness of his surroundings when a hand controller is used. For example, having to use up-down keys on the hand-held controller for menu navigation is much more demanding than speech-based control of these tasks.

5. Discussion and Recommendations

Even though the manual system was generally preferred, all of the participants said that the speech-based controller shows potential for robotic operation, assuming that improvements could be made to the voice control system. Potential areas for improvement include the elimination of the PTC button to minimize user errors, improving the vocabulary for better recognition, changing the pan and tilt command from a discrete to a continuous function, and adding macro commands to perform multiple actions. Additional studies may be required to evaluate any effect the modifications suggested in this report may have on task performance.

6. References

- Bravo, K. The Potential Reduction in Musculoskeletal Injury in the Nonscanning Arm by Using VoiceScan Technology During Sonographic Examinations. *Journal of Diagnostic Medical Sonography* **2005**, 21 (4), 304–308.
- Chang, E.; Meng, H.; Li, Y.; Fung, T. Efficient Web Search on Mobile Devices With Multi-Modal Input and Intelligent Text Summarization. *Proceedings of the Eleventh World Wide Web Conference*, Honolulu, HI, 7–11 May 2002.
- Draper, M.; Calhoun, G.; Ruff, H.; Williamson, D.; Barry, T. Manual vs. Speech Input for Unmanned Aerial Vehicle Control Station Operations. *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*; Human Factors and Ergonomics Society, Denver, CO, 2003, pp 109–113.
- Hart, S. G.; Staveland, L. E. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human Mental Workload*; Hancock, P. A., Meshkati, N., Eds.; Amsterdam, North Holland, 1988; pp 139–183.
- Mitchard, H.; Winkles, J. Experimental Comparisons of Data Entry by Automated Speech Recognition, Keyboard, and Mouse. *Human Factors* **2002**, 44 (2), 198–209.
- Pettitt, R.; Redden, E.; Carstens, C. *Scalability of Robotic Controllers: An Evaluation of Controller Options*; ARL-TR-4457; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2008.
- Pigeon, S.; Swail, C.; Geoffrois, E.; Bruckner, G.; Van Leeuwen, D.; Teixeira, C.; Orman, O.; Collins, P.; Anderson, T.; Grieco, J.; Zissman, M. *Use of Speech and Language Technology in Military Environments*; RTO-TR-IST-037; Research and Technology Organization of NATO, 2005, 3–11.
- Summers, G. *Evaluation of a Voice Activated Environmental Control System for Disabled People Clinical Rehabilitation* **1988**, 2 (3), 231–239.
- Tsimhoni, O.; Smith, D.; Green, P. Address Entry While Driving: Speech Recognition vs. a Touch Screen Keyboard. *Human Factors* **2004**, 46 (4), 600–610.

Turner, D. D.; Carstens, C. B. *Future Force Warrior: Insights From Air Assault Expeditionary Force Assessment*; ARL-TR-4191; U.S. Army Research Laboratory: Aberdeen Proving Ground, MD, 2007.

Williamson, D. T.; Barry, T. P.; Liggett, K. K. Flight Test Results of ITT VRS-1290 in NASA OV-10. *Proceedings of the 15th Annual International Voice Technologies Applications Conference AVIOS '96*; American Voice Input/Output Society, San Jose, CA, 1996; pp 335–345.

INTENTIONALLY LEFT BLANK.

Appendix A. Medical Status Form

This appendix appears in its original form, without editorial change.

Experiment participant: Please answer all questions honestly and completely. It will not be entered into your official health records and will be treated confidentially.

Roster Number: _____ Date: _____

1. Do you have any physical injury at the present time?

Yes _____ No _____

If yes, please describe. _____

2. Have you had any surgery in the last two months?

Yes _____ No _____

If yes, please describe. _____

3. Are you presently on a profile of any type? Yes _____ No _____

If yes, please describe your current limitations. _____

4. If the APFT (Army Physical Fitness Test) were held today, could you obtain a passing score on it? Yes _____ No _____

5. Do you have any medical concerns about carrying your combat fighting load while performing this exercise? Yes _____ No _____

If yes, please describe your concerns. _____

6. Have you had any type of eye surgery or eye injury? Yes _____ No _____

If yes, please describe. _____

Appendix B. Demographics

This appendix appears in its original form, without editorial change.

SAMPLE SIZE = 11

| <u>MOS</u> | <u>RANK</u> | <u>AGE</u> | <u>DUTY POSITION</u> |
|------------|-------------|-----------------|----------------------|
| O9S – 7 | E-4 – 5 | 30 years (mean) | OCS – 5 |
| 15P – 1 | E-5 – 2 | | NR – 6 |
| 44B – 1 | E-6 – 1 | | |
| 92W – 1 | NR – 3 | | |
| NR – 1 | | | |

1. How long have you served in the military? 55 months (mean)
2. How long have you been deployed overseas? 17 months (mean)
3. How long have you been deployed in a combat area? 10 months (mean)
4. What is your corrected visual acuity? 9 20/20 both eyes 1 20/20 one eye 1 NR
5. Please list any visual problems which you may have: Astigmatism (1).
6. What is your height? 69 inches (mean) (range 60-74)
7. What is your weight? 182 pounds (mean) (range 135-218)
8. With which hand do you most often write? 9 Right 2 Left
9. With which hand do you most often fire a weapon? 11 Right 0 Left
10. Do you wear prescription lenses? 2 Glasses 3 Contacts
11. Do you wear prescription lenses while firing a weapon? 6 No 5 Yes
12. Which is your dominant eye? 10 Right 1 Left
13. Have you ever driven a robotic vehicle? 11 No 0 Yes
14. Have you ever driven a remote control car? 1 No 10 Yes

15. Using the scale below, rate your level of experience with the following computer software and computer-related activities.

| | | | | | | |
|------------------|------------------|---------------------------|---------|---------------------------|------------------|--------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| No experience | Below average | Slightly below average | Average | Slightly above average | Above average | Expert |

| | MEAN RESPONSE |
|---|---------------|
| a. Microsoft Windows 98, 2000, XP, etc. | 5.27 |
| b. Computer-based games | 4.09 |
| c. Army digital systems (e.g., FBCB2) | 1.73 |
| d. I would self-rate my computer skills as: | 5.18 |

16. Using the scale below, please self-rate the following Knowledge, Skills, and Abilities (KSAs) related to infantry duties.

| | | | | | | |
|------------------|------------------|---------------------------|---------|---------------------------|------------------|--------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| No experience | Below average | Slightly below average | Average | Slightly above average | Above average | Expert |

| | MEAN RESPONSE |
|---|---------------|
| a. Knowledge of tactics, techniques, and procedures (TTP) | 2.55 |
| b. Knowledge of map reading and orientation in field setting | 3.82 |
| c. Knowledge of reconnaissance, surveillance, and target acquisition procedures | 2.27 |
| d. Knowledge relating to communications equipment and communications procedures | 2.45 |
| e. Communication skills (ability to use communications equipment and face-to-face communications to enhance mission accomplishment) | 4.18 |

INTENTIONALLY LEFT BLANK.

Appendix C. Training

This appendix appears in its original form, without editorial change.

TYPE/SAMPLE SIZE:

SPEECH-BASED = 11

MANUAL = 11

1. Using the scale below, please rate the training you received in the following areas.

1 2 3 4 5 6 7
Extremely bad Very bad Bad Neutral Good Very good Extremely good

| | MEAN RESPONSE | |
|--|---------------|--------|
| | Speech-based | Manual |
| Explanation of how to drive robot and operate controller | 6.55 | 6.55 |
| Time provided for explanation of the controller | 6.27 | 6.09 |
| Difficulty of controller operation | 4.91 | 5.40 |
| Time provided to practice driving the robot | 6.09 | 6.09 |
| Training on how to use this controller | 6.09 | 6.36 |
| Difficulty of learning the controller | 5.55 | 5.50 |
| Explanation of how to complete each station of course | 6.55 | 6.73 |
| Evaluation of the practice lane | 6.18 | 6.27 |
| How well you expect to perform on the actual course | 4.90 | 5.27 |
| Overall evaluation of the training course | 6.09 | 6.36 |

2. What were the easiest and hardest training tasks to learn?

Comments

No. of Responses

EASIEST

Speech-based

| | |
|--|---|
| Controlling arm by voice. | 1 |
| Voice accuracy. | 1 |
| Speech training. | 1 |
| Forward. | 1 |
| Halt, slow linear movements. | 1 |
| Learning commands. | 1 |
| Manual driving. | 1 |
| Operation of the controls. | 1 |
| Panning left and right. | 1 |
| What to do when, i.e., flippers down/arm high. | 1 |

Manual

| | |
|------------------------------------|---|
| All. | 1 |
| Basic front/back/turning movement. | 1 |

| <u>Comments</u> | <u>No. of Responses</u> |
|-------------------------------|--------------------------------|
| Controller. | 1 |
| Controlling the robot. | 1 |
| Drive forward. | 1 |
| Driving the machine. | 3 |
| Maneuvering. | 1 |
| Moving around with the robot. | 1 |
| Moving arm camera. | 1 |

HARDEST

| | |
|--------------------------------------|---|
| <u>Speech-based</u> | |
| Brake off. | 2 |
| Controlling zoom by voice. | 1 |
| Driving and speech at the same time. | 2 |
| Remembering commands. | 1 |
| Voice commands. | 1 |
| Execution at high rate of speed. | 1 |
| Maneuvering. | 1 |
| Releasing button. | 1 |

| | |
|--|---|
| <u>Manual</u> | |
| Arm movements requiring hitting buttons not on controller. | 1 |
| Camera pan/zoom. | 1 |
| Driving and maneuvering the camera. | 1 |
| Driving while giving verbal commands. | 1 |
| Flippers forward and backward. | 1 |
| Multitasking movements simultaneously. | 1 |
| Operating the arm. | 1 |
| Remembering commands, | 1 |
| Switching between joystick and screen. | 1 |
| Tracking target. | 1 |

3. What are your comments on the training course?

| | |
|--|---|
| <u>Speech-based</u> | |
| All in all, it was a great course. | 1 |
| Good instruction and training given by instructors. | 2 |
| It was a good way to test the use of the robot. | 1 |
| Very easy to navigate. | |
| I believe that a combination would be ideal. | 1 |
| Red flags hard to see. | 1 |
| Sun makes it hard to see engineers tape to stay in test area but that's a real world expectation; more familiarization with machine would make it easier to identify items that are out of the ordinary. | 1 |

| <u>Comments</u> | <u>No. of Responses</u> |
|--|--------------------------------|
| The course needs to be more on other terrains to simulate certain real world experience. Such as using flippers to go up hill and down hills, rocky roads, sandy roads, etc. | 1 |
| Training course requires a wide selection of commands. | 1 |
| <u>Manual</u> | |
| Good course. | 1 |
| Easy to learn manual controls, low learning curve. | 1 |
| Easy to navigate. | 1 |
| Use different terrain, features. Overall, it was great. | 1 |
| I feel the training course was designed to test many of the different ways to use the robot. | 1 |
| Red flags hard to see. | 1 |
| Could be more clearly marked (tape higher off ground?). | 1 |
| Sun made it difficult to see things. The simple tasks were harder than I thought they would be when you combine the two control types. | 1 |
| The camera should be operated with the right joystick. | 1 |
| 4. What are your comments on this controller? | |
| <u>Speech-based</u> | |
| Voice recognition overall is very good. | 1 |
| Easy to learn and easy to use. | 2 |
| The speech control makes a huge difference and with training will shorten mission times on the battlefield. | 1 |
| When speaking, it was easy. But adding in the manual and switching between the two when training was confusing on what I was supposed to be doing what with which type of command. If just speech, good; if speech and manual, need more training. | 1 |
| Giving voice recognition software the ability to continually learn a user's voice commands would be a great addition. | 1 |
| Doesn't always recognize command. | 1 |
| Having to push button all the time; seems cumbersome. | 1 |
| I did not like the operation of the "push to talk" button. | 1 |
| It's hard to remember voice used when presetting the controller. People's tones change throughout the day and I would have to uphold the same tone, or the robot would not respond. | 1 |
| So much harder to control movements with voice commands compared to manual. Fine movements much more difficult. Voice recognition of commands was extremely high. A synergy of controls would prove superior, especially with "arm drive" command and other arm commands-- not drive commands. | 1 |

Comments**No. of Responses****Manual**

| | |
|--|---|
| I liked the controller used to navigate the robot. It was easy to learn and easy to use. | 1 |
| Accurate and easy controls, except for tracking moving target - especially when zoomed in. | 1 |
| Easy to use, high level of command after very little practice. Worked better with some combination of voice to manual (halt, arm drive, a few others). | 1 |
| I'm not a gamer, so for me to remember the controls was difficult, but easy to figure out. | 1 |
| Manual operation is difficult to combine movements, but with continued training, wouldn't be a problem. | 1 |
| A lot of buttons to remember. | 1 |
| Camera control on opposite side of drive controls. | 1 |
| I think moving the camera functions to the right joystick would allow the keypad to be used for the arm function, which would save the operator from having to press buttons on the computer. | 1 |
| Integrating screen commands to hand controller. | 1 |
| Not a video gamer, so all the buttons are not intuitive to me. Maybe a track ball of some kind? But with more training, would prefer controller to speech. More flipper controls to left hand so drive movements are on the same hand. Having to reach the screen while still maintaining the controller is difficult. | 1 |

INTENTIONALLY LEFT BLANK.

Appendix D. Post Iteration

This appendix appears in its original form, without editorial change.

TYPE/SAMPLE SIZE:**SPEECH-BASED = 11****MANUAL = 11**

1. Using the scale below, please rate your ability to perform each of the following control tasks based on your experience with the controller you used.

1 2 3 4 5 6 7
Extremely difficult Very difficult Difficult Neutral Easy Very easy Extremely easy

| | MEAN RESPONSE | |
|---|---------------|--------|
| | Speech-based | Manual |
| Position robot to look through a window | 6.00 | 6.36 |
| Position robot to look over a wall | 6.09 | 6.45 |
| Scan for targets | 5.00 | 6.36 |
| Track a target | 4.27 | 6.18 |
| Zoom in and out with the camera | 4.67 | 6.27 |
| Maintain driving control while negotiating the course | 5.73 | 6.27 |
| Position robot to look in a bunker | 5.55 | 5.55 |
| Drive into a tunnel | 6.45 | 6.64 |
| Stay within course limits (engineer tape) | 5.91 | 6.27 |
| Control robot's rate of speed | 5.73 | 6.09 |
| Raise robot arm | 6.36 | 5.27 |
| Tilt sensor head | 5.64 | 6.30 |
| Pan sensor head | 5.90 | 6.27 |
| Steer robot while moving forward | 5.64 | 6.45 |
| Steer robot while moving in reverse | 5.60 | 6.09 |
| Position robot flippers | 6.45 | 6.00 |
| Perform multiple tasks simultaneously (i.e., raise arm while driving) | 5.45 | 4.55 |
| Ease of getting the speech-based system to understand your commands | 5.18 | 6.33 |
| Overall ability to perform driving tasks | 5.36 | 6.18 |

Comments**No. of Responses****Speech-based**

Voice commands make control easier. 1
 More difficult to control in reverse. Have to reposition drive camera, arm camera, 1
 and perspective not as good.

Comments**No. of Responses**

Robot didn't perform the correct commands. 1

Some of the verbal commands did not register until the 2nd or 3rd time. 1

Unit is sometimes not responsive to voice commands and must be repeated several times. 1

Speech recognition very hard. Confuses commands. 1

When covering large areas, hand controller is better. But on a hand controller maybe have a hold down button to use while moving that will slow down the driving. 1

Zoom - there was only one speed. When you zoom, it would go too far. 1

Manual

Easy to pan and track a target. 1

Very easy to control with game pad. 1

Arm too low to look in a bunker with slightly raised sand bags. Maybe need a medium height arm location. 1

Difficult to punch buttons on console while driving with control. 1

Panning the camera while driving is awkward having to reach across with my right hand. 1

Switching between controller and screen is awkward, especially while maintaining movement. 1

2. During this trial, did you experience any of the following problems?

| | % of YES Responses | |
|----------------|--------------------|--------|
| | Speech-based | Manual |
| Wrist strain | - | - |
| Arm strain | - | - |
| Hand strain | - | - |
| Finger strain | 9% | - |
| Nausea | - | - |
| Fatigue | - | - |
| Disorientation | 9% | 18% |
| Dizziness | - | - |
| Other | - | - |

Speech-based

Current button made my finger hurt. 1

Small camera view posed an orientation problem for me. 1

Manual

It was difficult maintaining orientation with the small screen/view. 1

Remembering which camera to look in for driving vs. when using the camera specific and using both at same time when approaching a target. 1

3. Using the scale below, please rate the following characteristics of the controller that you used.

1 2 3 4 5 6 7
Extremely bad Very bad Bad Neutral Good Very good Extremely good

| | MEAN RESPONSE | |
|---------------------------------|---------------|--------|
| | Speech-based | Manual |
| Control responsiveness | 5.73 | 5.64 |
| Control sensitivity | 5.27 | 5.91 |
| Location of buttons | 6.11 | 5.73 |
| Mapping of controller functions | 6.33 | 6.00 |

Comments

No. of Responses

Speech-based

It is a good design to navigate the robot; easy to learn and use. 1

Response was slow on panning. 1

Manual

Controller is easy to use and navigate the robot. 2

Didn't like having to use the buttons on the monitor. 1

Location of buttons OK due to being a general remote. Something built specific for might work better. Move flipper buttons to left side of controller so all drive functions are on the same side. 1

Panning the camera while driving is awkward having to reach across with my right hand. 1

Raising arm buttons are on a console. 1

4. Using the scale below, what is your overall rating of the controller that you used?

1 2 3 4 5 6 7
Extremely bad Very bad Bad Neutral Good Very good Extremely good

| MEAN RESPONSE | |
|---------------|--------|
| Speech-based | Manual |
| 5.40 | 5.91 |

Speech-based

Controller performed well. 1

Controller is good. 1

Intuitive controls/buttons on controller and good layout of PCC functional/ menu buttons. Fairly easy to control with little training for basic functions. 1

| <u>Comments</u> | <u>No. of Responses</u> |
|---|--------------------------------|
| Over a large area, hand control better because easier to have a conversation/radio transmission. For me that is an easier way to multitask. Hands do one thing and can have conversation about something else. | 1 |
| Verbal commands were easier for everything except driving because I was able to control driving much better with controls. | 1 |
| Pressure sensitivity helped control the speed of the robot. | 1 |
| Perhaps removing any buttons that won't get used to prevent confusion during combat use. | 1 |
| I've noticed when giving speech commands and someone is talking behind me, the voice commands were not registering correctly. At that moment I wanted to revert to controller, but had to use voice commands. | 1 |
| Speech-based has problems when recognizing commands. | 1 |
| Tilting/panning camera relatively difficult and jerky. | 1 |
| <u>Manual</u> | |
| The controller is good. | 1 |
| This course was fun. I felt more at ease this time around with practice. | 1 |
| Very easy to use and ergonomic. | 1 |
| Camera pan/tilt should be on right joystick and arm controls should be moved to keypad on controller. | 1 |
| Controller being not specific to the task and having to reach forward to move arm not so good. Maybe arm control buttons on the right side underneath the arm like maybe on the arm rest. | 1 |
| Controller's buttons are well positioned to help navigate; however, reaching over to move the robots arm is awkward and makes you lose focus a little on what you are doing. Overall, I prefer the manual over the speech but in some instances, it is easier using speech commands. For example, on the 3rd test, where you have to move thru the grass locating IEDs, it is easier using speech commands. But when the camera is looking thru the window and tilting, it is easier to use the manual controller. I also found it easier to use manual commands while tracking the target. | 1 |
| Only hard part is manipulating arm with buttons not located on controller (raise arm, arm high, etc.) while attempting to maintain motion (simultaneous tasks). | 1 |

INTENTIONALLY LEFT BLANK.

Appendix E. End of Experiment

This appendix appears in its original form, without editorial change.

SAMPLE SIZE = 11

1. Please rank order the controller condition in the order of your preference.

| Controller Condition | # preferred | % preferred |
|-----------------------------|--------------------|--------------------|
| Manual | 7 | 64% |
| Speech-based | 3 | 27% |
| No response | 1 | 9% |
| Total | 11 | 100% |

Comments

No. of Responses

| | |
|---|---|
| A combination of the two would be my favorite. | 4 |
| I preferred maneuvering with manual while deploying the arm with voice. | 1 |
| Manual controls felt far more natural, greater fire control. | 1 |
| Manual easy to learn. | 1 |
| Manual was easier and worked every time. | 1 |
| Speech-based was much easier than manual simply because giving a command and then moving to the next movement knowing that your voice command is being executed makes controlling easier. | 1 |
| Favored the speech-based because less effort. | 1 |
| Speech-based was a lot easier to manipulate things at once. | 1 |
| Speech recognition is very good, but even a rare glitch can cause the vehicle to perform a command not intended, resulting in excessive time delays or corrections. | 1 |
| Multitasking with voice when you have radio traffic and other things you need to listen for is difficult. | 1 |
| The voice commands were frustrating, i.e., if you did not push the button long enough, etc. | 1 |
| Speech-based had many bugs in system. | 1 |

2. Do you have suggestions for ways to increase the effectiveness of the controllers?

Manual

| | |
|--|---|
| Add screen controls to handheld. | 1 |
| Better locations of buttons. | 1 |
| Use all the buttons, e.g., button one could be arm high/low with shift key instead of no function. | 1 |
| The green button used for voice needs to be mounted on controller, not Velcroed on. I found it to be very uncomfortable to drive and speak at same time. | 1 |

Comments**No. of Responses**

| | |
|--|---|
| Cushioning on sides. | 1 |
| Dedicated controller for the task. For arm deployment, a controller to put under hand as if on hand rest to control drive arm. | 1 |
| Moving the pan/tilt to the right joystick. | 2 |
| Smoother movement while tracking a moving target with the pan function. | 1 |
| Wider angle camera on drive camera. | 1 |

Speech-based

| | |
|---|---|
| Occasional command unrecognized, but overall very good. | 1 |
| A lot of bugs that confuse commands. | 1 |
| Jerky, strong voice recognition. Need to be able to pan/tilt more than just set increments. | 1 |
| Panning should be continuous until I say stop, instead of me repeating myself. | 2 |
| Replace push to talk button with a voice recognition start command, incremental panning by degrees. | 1 |
| Use a headset rather than a push button. | 1 |
| More immediate response. | 1 |
| Also, instead of preset pan/tilt movements, make it continuous with a "stop" command. | 1 |

3. Do you have any suggestions for alternative words to be used for the following speech-based controller commands?

| | |
|---------------------------|---|
| Halt | No change (7), stop (3), don't have to say "brake off " afterwards |
| Forward large and small | No change (4), forward fast & slow (6), strong & weak |
| Reverse large and small | No change (5), reverse fast & slow (5), strong & weak |
| Left large and small | No change (6), fast & slow (3), strong & weak, left medium |
| Right large and small | No change (6), fast & slow (3), strong & weak, right medium |
| Pan left and right | No change (8), scan left & right, pan slow left & right, continuous instead of repeating myself |
| Tilt up and down | No change (9), pan up & down, continuous instead of repeating myself |
| Zoom in and out | No change (10), continuous instead of repeating myself |
| Lights on and off | No change (11) |
| Flippers forward and back | No change (8), up & down (3) |
| Brake on and off | No change (10), get rid of |
| Arm drive | No change (8), reset (2), arm down |
| Arm deploy | No change (8), arm mid (2), extend |
| Arm retract | No change (11) |
| Arm high | No change (11) |
| Arm low | No change (11) |

* (#) denotes number of responses

Comments**No. of Responses**

Commands well understood by software.

1

4. Do you think that a speech-based controller shows potential for robotic operation?

11 Yes 0 No

Being able to give voice commands for flippers and arm movement while approaching an objective would be better than a manual movement because you can pay more attention to your objective. 1

Easier than reaching up and pressing buttons on screen. 1

Further testing and more time training will give any operator better results. 1

If voice recognition could be improved, I think that speech-based controllers show a lot of potential. 1

It's very good. I just find it frustrating if in a noisy environment, the voice command does not pick up sounds, but assumes commands. 1

Some form of combination seems best. 1

5. Using the scale below, please rate the importance of the following information that is needed for teleoperating a robot.

| | | | | | | |
|-------------|-------------|-------------|---------|-----------|-----------|-----------|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| Extremely | Very | Unimportant | Neutral | Important | Very | Extremely |
| unimportant | unimportant | | | | important | important |

| DRIVING TASKS | MEAN RESPONSE |
|--|----------------------|
| Heading information. | 5.56 |
| Relative distance to obstacles | 6.40 |
| Depth of pot holes | 5.50 |
| Information on the navigatability of down slopes (tilt meter) | 5.60 |
| Information on the navigatability of side slopes (tilt meter) | 5.70 |
| Identification of any other terrain features that might have an adverse effect on the ability of the robot to maneuver through the terrain | 5.55 |
| Ability to determine where the sides of the vehicle are located (sides are in the camera field of view or are marked on the screen) | 5.50 |
| Ability to determine where the front of the vehicle is located | 5.70 |
| Ability to determine where the back of the vehicle is located | 5.20 |
| Information concerning whether the ground clearance of the vehicle will allow negotiation of rugged terrain | 5.70 |
| Indication of the turn radius of the vehicle | 5.40 |
| Information on the terrain closer than 6 inches in front of the vehicle | 5.20 |
| Information on the terrain five to 15 feet in front of the vehicle | 5.80 |
| Information on the terrain greater than 15 feet in front of the vehicle | 5.20 |
| Information on the current speed of the vehicle (speedometer) | 4.40 |

(cont)

| DRIVING TASKS | MEAN RESPONSE |
|---|--------------------------|
| Information on vehicle RPMs | 3.30 |
| Information on the way the vehicle motor sounds | 4.00 |
| Information on the other noises present in the vehicle's environment | 4.70 |
| Information concerning the color of objects in the environment (color vs black and white camera) | 6.18 |
| Information concerning objects on the side of the vehicle (side facing camera) | 4.50 |
| Information concerning objects behind the vehicle (back facing camera) | 5.40 |
| Information concerning the condition of the vehicle tires/tracks | 4.60 |
| Information concerning the temperature of the vehicle | 4.90 |
| Feedback on the ruggedness of terrain | 4.30 |
| Information concerning any delay between the time the operator sends a vehicle command and the time that the vehicle responds (i.e., how long after a stop command is sent before the vehicle actually stops) | 6.00 |
| Feedback concerning whether or not the vehicle correctly received a command | 6.20 |
| Battery status | 6.55 |
| Fuel status | 6.60 |
| Vehicle temperature | 5.50 |

Comments

No. of Responses

Information provided seemed sufficient. Temperature seems very irrelevant info; same with ruggedness of terrain. I don't care about vehicle's RPMs or the noise it makes. I want to see the pothole so I can avoid it, not determine its depth.

1

NO. OF
COPIES ORGANIZATION

1 DEFENSE TECHNICAL
(PDF INFORMATION CTR
only) DTIC OCA
8725 JOHN J KINGMAN RD
STE 0944
FORT BELVOIR VA 22060-6218

1 DIRECTOR
US ARMY RESEARCH LAB
IMNE ALC HRR
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 DIRECTOR
US ARMY RESEARCH LAB
RDRL CIM L
2800 POWDER MILL RD
ADELPHI MD 20783-1197

1 DIRECTOR
US ARMY RESEARCH LAB
RDRL CIM P
2800 POWDER MILL RD
ADELPHI MD 20783-1197

ABERDEEN PROVING GROUND

1 DIR USARL
RDRL CIM G (BLDG 4600)

NO. OF
COPIES ORGANIZATION

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR ML J MARTIN
MYER CENTER RM 2D311
FT MONMOUTH NJ 07703-5601

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MZ A DAVISON
320 MANSCEN LOOP STE 115
FT LEONARD WOOD MO 65473

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MD T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL 35898-7290

1 COMMANDANT USAADASCH
ATSA CD
AMSRD ARL HR ME DR HAWLEY
5800 CARTER RD
FT BLISS TX 79916-3802

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MM DR V J RICE
BLDG 4011 RM 217
1750 GREELEY RD
FT SAM HOUSTON TX 78234-5002

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MG R SPINE
BLDG 333
PICATINNY ARSENAL NJ 07806-5000

1 ARMY RSCH LABORATORY - HRED
ARMC FIELD ELEMENT
AMSRD ARL HR MH C BURNS
THIRD AVE BLDG 1467B RM 336
FT KNOX KY 40121

1 ARMY RSCH LABORATORY - HRED
AWC FIELD ELEMENT
AMSRD ARL HR MJ D DURBIN
BLDG 4506 (DCD) RM 107
FT RUCKER AL 36362-5000

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MK MR J REINHART
10125 KINGMAN RD
FT BELVOIR VA 22060-5828

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MY M BARNES
2520 HEALY AVE
STE 1172 BLDG 51005
FT HUACHUCA AZ 85613-7069

NO. OF
COPIES ORGANIZATION

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MP D UNGVARSKY
POPE HALL BLDG 4709
BCBL 806 HARRISON DR
FT LEAVENWORTH KS 66027-2302

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MJF J HANSBERGER
JFCOM JOINT EXPERIMENTATION J9
JOINT FUTURES LAB
115 LAKEVIEW PKWY STE B
SUFFOLK VA 23435

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MQ M R FLETCHER
NATICK SOLDIER CTR
AMSRD NSC WS E BLDG 3 RM 343
NATICK MA 01760-5020

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MT J CHEN
12423 RESEARCH PKWY
ORLANDO FL 32826

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MT C KORTENHAUS
12350 RESEARCH PKWY
ORLANDO FL 32826

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MS C MANASCO
SIGNAL TOWERS
BLDG 29808A RM 303
FT GORDON GA 30905-5233

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MU
6501 E 11 MILE RD MS 284
BLDG 200A 2ND FL RM 2104
WARREN MI 48397-5000

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MF C HERNANDEZ
2421 NW AUSTIN RD STE 220
FT SILL OK 73503-9043

1 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MV
S MIDDLEBROOKS
91012 STATION AVE RM 348
FT HOOD TX 76544-5073

NO. OF
COPIES ORGANIZATION

- | | |
|-------------------|---|
| 1 | ARMY RSCH LABORATORY - HRED AMSRD ARL HR MN R SPENCER DCSFDI HF HQ USASOC BLDG E2929 FT BRAGG NC 28310-5000 |
| 1 | ARMY RSCH LABORATORY - HRED AMSRD ARL HR MW E REDDEN BLDG 4 ROOM 332 FT BENNING GA 31905-5400 |
| 1 (CD only) | ARMY G1 DAPE MR B KNAPP 300 ARMY PENTAGON RM 2C489 WASHINGTON DC 20310-0300 |

ABERDEEN PROVING GROUND

- | | |
|---|---|
| 7 | DIR USARL RDRL CIM G S FOPPIANO RDRL HR L ALLENDER P FEDELE T LETOWSKI RDRL HRM B J LOCKETT RDRL HRM D D HARRAH RDRL HRS D B AMREIN |
|---|---|

NO. OF
COPIES ORGANIZATION

30 ARMY RSCH LABORATORY - HRED
AMSRD ARL HR MW
C CARSTENS (10 CPS)
R PETTITT (10 CPS)
E REDDEN (10 CPS)
BLDG 4 RM 60
FORT BENNING GA 31905

INTENTIONALLY LEFT BLANK.